arXiv:1203.0313v1 [physics.soc-ph] 1 Mar 2012

Statistical Analysis of the Road Network of India

Satyam Mukherjee

Department of Chemical and Biological Engineering, Northwestern University Evanston, Illinois - 60208

Abstract. In this paper we study the Indian Highway Network as a complex network where the junction points are considered as nodes, and the links are formed by an existing connection. We explore the topological properties and community structure of the network. We observe that the Indian Highway Network displays small world properties and is assortative in nature. We also identify the most important road-junctions (or cities) in the highway network based on the betweenness centrality of the node. This could help in identifying the potential congestion points in the network. Our study is of practical importance and could provide a novel approach to reduce congestion and improve the performance of the highway network.

1. Introduction

Transportation networks form the backbone of economic development in a country. In the recent years, tools from network analysis have been applied to several transportation networks around the world. These include the airport network of China [1], the airport network of India [2], the world-wide airport network [3,4], the urban road networks [5] and the railway networks [6–10]. The topology studies of different spatial networks show different degree distributions. Power law degree distribution is seen for Indian airport network [2] and world-wide airport network [3,4]. Also, two regime power laws is observed in China airport network [1] and US airport network [11]. On the other hand exponential degree distribution was observed in railway networks of India [7] and China [8].

It is to be noted that several models have been proposed to study various transportation networks. The observed topological properties depend on the way the network is modeled. A directed network model was used to study the Chinese railway network [8]. Many transportation networks have been modeled as bipartite networks and weighted networks [9,12]. Again, urban networks have been modeled on local optimization process [13].

Road networks are one of the most prominent means of transport in many countries around the globe and display complex topological properties. For example, urban road networks in Le Mans (France) show double-power law degree distribution [14]. The Indian Highway Network (IHN) is one of the busiest road networks in the world, constituting of 2% of all roads in India but handling 40% of the total road traffic as of 2010 [15]. The total length of the highways extend up to $70,934$ km. National highways are the backbone
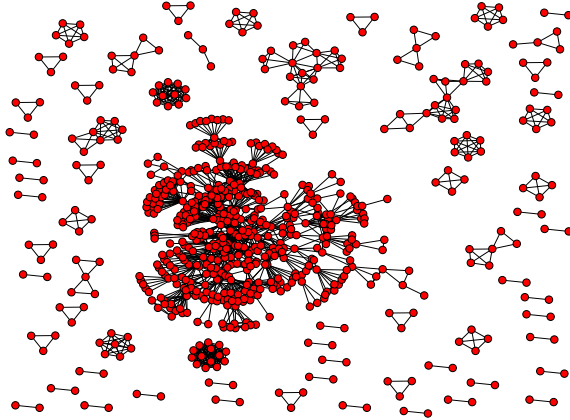
Figure 1. (Colour online) Network structure of Indian highways.

of economic development in the country as many cities, towns and industries have come up along the highways. However, according to recent surveys, India faces the challenge of poor rural roads and traffic congestion in major cities. India's road network logistics and bottlenecks in transportation hinder its GDP 1% to 2% (equivalent to US$16 billion - US$32 billion). In this situation it is therefore important to study the topological features of the highway network. Such a study could help in designing efficient modes of extension of the IHN and also a better planning of handling traffic congestion. In this paper we study the Indian highway network as an undirected network. Two nodes(cities/junctions) are connected if there exists a common bus route between them. We measure the basic topological properties of the present IHN and identify potential points of congestion in the network.

The rest of the paper is laid out as follows. In Section 2 we discuss the data and method of network construction and the topological properties of the network. In Section 3 we have described the community analysis of the Indian highway network. We conclude in Section 4.

2. Network construction and Topological Properties

The highway network is generated as follows. Two cities b_i and b_j are connected if there exist at least one bus-route which links the two stations. Such a representation has been used earlier in the context of other transportation networks like the railway network [7,10] and Indian airport network [2]. Figure 1 shows the Indian Highway Network (IHN). From the figure it is evident that the IHN has a giant component (which is the largest connected component of the network) and nodes which are not connected to the giant component. This indicates that the IHN is not a fully conneced network. The IHN consist of 694 nodes and 2209 edges. We analyse the topological properties of the IHN. Clustering coefficient (C_i) of a node i is defined as the ratio of number of links shared by its neighboring nodes to the maximum number of possible links among them. The average clustering coefficient
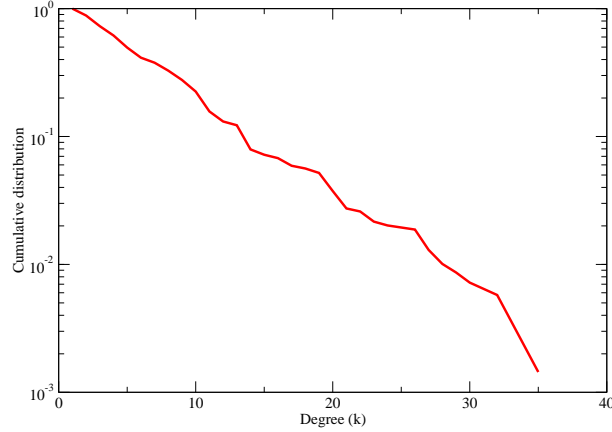
2

Figure 2. (Colour online) Cumulative degree distribution of the IHN.

is defined as,

$$C = \frac{1}{N} \sum_{i=0}^{N} C_i \qquad (1)$$

We observe that the average clustering coefficient (C) of the network is 0.78, indicating that the IHN is a highly clustered network. In Table 1 we compare the topological properties of IHN with Erdös-Rényi model (ER model) [21] and Configuration model [22]. Keeping the network size same as that of IHN, the network based on configuration model is simulated considering the same degree sequence of IHN. We observe that the average clustering coefficient (C) of the network is 0.78, indicating that the IHN is a highly clustered network. The clustering coefficient of ER random graph (0.008) is an order of magnitude lower than that of the IHN (See Table 1). We also determine the average shortest path length D between a given vertex and all other vertices of the IHN. The average shortest path length is defined as :

$$D = \sum_{s,t \epsilon V} \frac{d(s,t)}{N(N-1)} \qquad (2)$$

where V is the set of nodes in the network, $d(s,t)$ is the shortest path from s to t, and N is the number of nodes in the network. As shown in Table 1 we observe that D for IHN is of the same order of magnitude as that of ER random graph. These two properties indicate that the IHN is a small world network.

The degree distribution of a network is defined as the fraction of nodes with degree k in the network. In the case of IHN the degree is defined as the number of cities (junctions) that can be reached from a given city (junction) via a single direct bus. In Figure 2 we show the cumulative degree distribution of the IHN. Another parameter which is studied is the average degree of nearest neighbours, $K_{nn}(k)$ for nodes of degree k (Figure 3). We observe that $K_{nn}(k)$ increases with degree. This indicates that strong correlations are present among nodes of different degrees. This behaviour is contrary to that observed in the Indian Railway Network [7]. The topological assortativity coefficient ($-1 \leq r \leq$

3

Table 1. Comparison of network properties of Indian Highway Network (IHN) with Erdös-Rényi model ($p = 0.009186$) and Configuration model.

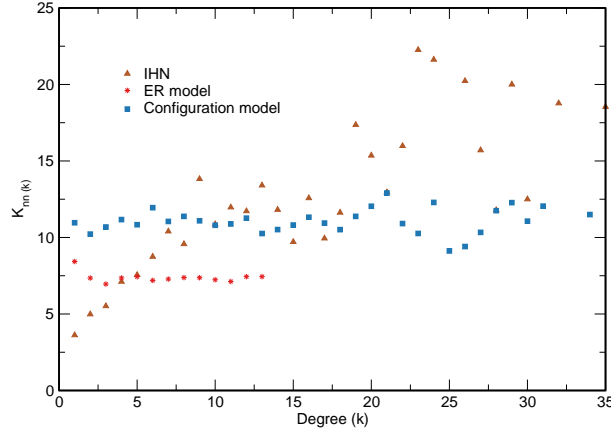| $Properties$	IHN	$Erdös - Rényi$	$Configuration$
v	694	694	694
e	2209	2148	2169
$\bar{k}$	6.36	6.19	6.25
C	0.78	0.008	0.015
D	4.75	3.77	3.61

Figure 3. (Colour online) Average unweighted ($K_{nn}(k)$) degree of nearest neighbours of nodes with degree k.

1), as defined by Newman is another measure of degree correlations in a network [16]. Mathematically, its defined as

$$r = \frac{1}{\sigma_q^2} \sum_{jq} jk(e_{jk} - q_j q_k) \tag{3}$$

where $q_k = \sum_j e_{jk}$ and $\sigma_q^2 = \sum_k k^2 q_k - \sum_k k q_k{}^2$. Here e_{jk} is the probability that a randomly chosen edge has nodes with degree j and k at either end. If the networks are of assortative nature, high degree nodes at one end of a link show some preference towards high degree nodes at the other end. For networks with disassortative mixing, high degree nodes at one end link show preference towards low degree nodes at the other end and vice versa. For the IHN $r = 0.46$, indicating a assortative mixing which is consistent with the observation in Figure 3.

We also identify the major cities in the IHN. The nodes with high betweenness centrality are potential points of traffic congestion in networks [17,18]. Betweenness centrality of a node v is the sum of the fraction of all-pairs shortest paths that pass through v:

$$c_B(v) = \sum_{s,t \epsilon V} \frac{\sigma(s,t|v)}{\sigma(s,t)} \tag{4}$$

4

Table 2. Top 10 cities of the IHN based on the betweenness centrality (BC).

BC	$Junction$
0.110	Delhi
0.107	Varanasi
0.081	Bangalore
0.051	Lucknow
0.050	Muzaffarpur
0.040	Ranchi
0.03837	Raipur
0.03833	Mangalore
0.0375	Jalandhar
0.0369	Chennai

where V is the set of nodes, $\sigma(s, t)$ is the total number of shortest paths and $\sigma(s, t|v)$ is the number os hortest paths passing through v [19]. We list the top ten cities based on the betweenness centrality (See Table 2). We identify *Delhi* as top city according to ranking based on betweenness centrality. This could be *posteriori* justified given the fact that *Delhi* is the capital city of India and handles maximum traffic.

3. Community analysis

We perform a community analysis on the giant-component of the IHN. To identify the communities in the IHN, we used the definition of modularity introduced by Newman [20]. As described in [20], modularity is defined as the number of edges falling within groups minus the expected number of edges in an equal sized network of randomly placed edges. Positive values of modularity Q, indicates possible presence of community structure. Here we optimize the modularity over possible divisions of the IHN and search for community structure of IHN by looking for positive and large values of Q. In Figure 4 we display the communities identified by the Newman algorithm in the IHN. We observe that *Delhi*, *Allahabad*, *Lucknow*, *Kanpur*, *Mathura*, *Ludhiana*, *Faizabad* and *Agra* share the same community. This is evident from the fact that all these cities are located in the Northern parts of India. Interestingly *Mumbai* and *Dankuni* share the same community with *Delhi* even though these two cities are located in the Western and Eastern region of India respectively. Similarly *Varanasi* is grouped with the Southern cities like *Bangalore*, *Mysore*, *Hyderabad*, *Cochin*, *Coimbatore* and *Kanyakumari*. This is also justified by the fact that the longest highway in India is *NH7* which stretches from *Varansi* in north India to *Kanyakumari* in the southern most point of India. We also observe that *Chennai* is grouped with *Pune*, *Belgaum* and *Bijapur*.

4. Conclusion

In this paper we have studied the Indian Highway Network (IHN) as an unweighted graph of bus-stations. We observed that the distribution of node-connectivity is neither a power-law nor normal. The IHN shows small world properties and assortative mixing. The cities (junctions) with high betweenness centrality have also been identified. We believe these
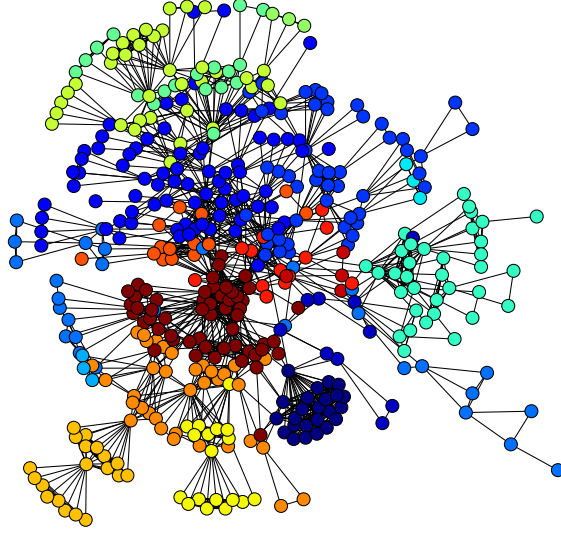
Figure 4. (Colour online) Each node indicates a city and each colour corresponds to a community. 17 communities are identified. Modularity $Q = 0.7718$.

nodes (cities/junctions) are potential points of congestion. Considering the limited capacity of links to handle road traffic, identification of possible congestion points is of practical importance. With the availability of traffic data, it would also be interesting to study the weighted IHN as it could reveal a clearer picture of network dynamics and the role of cities within a community. The traffic handling capacities of important cities and also that of the links should be enhanced and infrastructure improved in order to achieve efficient traffic. The IHN is expected to grow in size and achieve maximum connectivity. It would be interesting to study the evolution of the IHN in order to model the topology of the network. Given the availability of traffic data in different cities one could also implement the gravity model on Indian highways as was done for Korean highways [23]. It would also be interesting to study the economic growth along the highways and how it is correlated with the centrality of the cities and important junctions. Another possible direction of research is to analyze the road network of metropolitan cities and rural regions, which currently we are unable to study due to unavailability of data.

ACKNOWLEDGMENTS

References

[1] W. Li and X. Cai, Phys. Rev. E **69**, 046106 (2004)

[2] G. Bagler, Physica A **387**, 2972 (2008)

[3] R. Guimera, S. Mossa, A. Turtschi, L.A.N. Amaral, Proc. Natl. Acad. Sci. USA **102** 7794 (2005)

[4] A. Barrat, M. Barthlemy, R. Pastor-Satorras, A. Vespignani, Proc. Natl. Acad. Sci. USA **101**, 3747 (2004)

[5] J. Sienkiewicz, J.A. Holyst, Phys. Rev. E **72** 046127 (2005)

[6] K. S. Kim, L. Benguigui, M. Marinov, Cities **20** 31 (2003)

[7] P. Sen et al., Phys. Rev. E **67** 036106 (2003)

[8] W. Li and X. Cai, Physica A **382**, 693 (2007)

[9] K. Seaton, L. Hacket, Physica A **339**, 635 (2004)

[10] S. Ghosh, S. Banerjee, N. Sharma, S. Agarwal, N. Ganguly, S. Bhattacharya, A. Mukherjee, Acta Physica Polonica B Proceedings Supplement **4**, 123-137 (2011)

[11] C. Li-Ping et al., Chinese Phys. Lett. **20**, 1393 (2003)

[12] Y. Wang et al., Physica A **388**, 2949 (2009)

[13] Marc Barthélemy and Alessandro Flammini, Phys. Rev. Lett. **13** 138702 (2008)

[14] J. Jiang, Frederic Metz, Christophe Beck, Sebastien Lefebre, Jinchan Chen, Qiuping A. Wang and Michel Pezeril, International Journal of Modern Physics C **22** 13-20 (2011)

[15] National Highway Authority of India (*www.nhai.org*)

[16] M. E. J. Newman, *SIAM Review* 45, 167 (2003)

[17] B. K. Singh and N. Gupte, Phys. Rev. E **71**, 055103(R) (2005)

[18] S. Mukherjee and N. Gupte, Phys. Rev. E. **77**, 036121 (2008)

[19] U. Brandes, Social Networks **30**(2), 136 (2008)

[20] M. E. J. Newman, Proceedings of the National Academy of Sciences **103**(23) 8577 (2006)

[21] P. Erdös and A. Rényi. *On the evolution of random graphs*. Math. Inst. Hung. Acad. Sci. **5**, 7 (1960)

[22] M. E. J. Newman, S. H. Strogatz and D. J. Watts, Phys. Rev. E. **64**, 026118 (2001)

[23] Woo-Sung Jung, Fengzhong Wang and H. Eugene Stanley, Euro. Phys. Lett. **81** 48005 (2008)